

A novelty detection approach to diagnosing hull and propeller fouling

Andrea Coraddu^{a,*}, Serena Lim^c, Luca Oneto^b, Kayvan Pazouki^c, Rose Norman^c, Alan John Murphy^c

^a*Naval Architecture, Ocean & Marine Engineering
Strathclyde University, Glasgow G1 1XW, UK*

^b*DIBRIS - University of Genova, Via Opera Pia 13, I-16145 Genova, Italy*

^c*Marine, Offshore and Subsea Technology Group, School of Engineering
Newcastle University, Newcastle upon Tyne, NE1 7RU, UK*

Abstract

Hull and propeller performance have a primary role in overall vessel efficiency. Vessel fouling is a common phenomenon where undesirable substances attach or grow on the ship hull. A clear understanding of the extent of the degradation of the hull will allow better management of assets and prediction of the best time for dry docking and hull maintenance work. In this paper, the authors investigate the problems of predicting the hull condition in real operations based on data measured by the on-board systems. The proposed solution uses an unsupervised Machine Learning (ML) modelling technique to eliminate the need for collecting labeled data related to the hull and propeller fouling condition. Two anomaly detection methods based on Support Vector Machines and k-nearest neighbour have been applied to predict the hull condition using the available parameters measured on-board. Data from the Research Vessel The Princess Royal has been exploited to show the effectiveness of the proposed methods and to benchmark them in a realistic maritime application.

Keywords: Sensor Data Collection, Ship Efficiency, Hull and Propeller

*Corresponding Author

Email addresses: andrea.coraddu@strath.ac.uk (Andrea Coraddu), serena.lim@newcastle.ac.uk (Serena Lim), luca.oneto@unige.it (Luca Oneto), kayvan.pazouki@newcastle.ac.uk (Kayvan Pazouki), rose.norman@newcastle.ac.uk (Rose Norman), a.j.murphy@newcastle.ac.uk (Alan John Murphy)

1. Introduction

Hull and propeller performance have a pivotal role in ship efficiency. According to the Clean Shipping Coalition, poor hull and propeller performance are responsible for around 10% of world fleet energy costs and Greenhouse Gases emissions (IMO, 2011). This data is translated into billions of dollars in extra cost per year and contributes to a 0.3% increase in Greenhouse Gas emissions. This is caused by a combination of mechanical damage and biofouling.

Ship hulls are designed to withstand hydrostatic and hydrodynamic loading while operating at designed speed (Tupper, 2013). The hull design is streamlined to minimise the resistance when the body moves through the water, however the surface condition of the hull deteriorates over time due to fouling (Soares et al., 2009). Hull degradation may be caused by ship fouling, which is the common phenomenon where living organisms such as barnacles, slime and seaweed grow on the surface of the hull (Flemming, 2002). This creates additional viscous resistance for the vessel during operations (Schultz, 2007). The difference in applied coatings, type of material of the vessel hull, and different operating surroundings causes the hull condition to degrade at different levels and intensities (Kane, 2012). A clear understanding of the extent of degradation of the hull will allow better management of assets and the prediction of the best time for dry docking and hull maintenance work (Schultz et al., 2011).

Hull condition assessment is carried out typically through surveys and routine dry docking. The typical approach to assess the hull condition is to compare vessel performance to the speed performance curve. This curve is developed from the ship log and sea trials at specific loading conditions. The deviation from the speed performance curve cannot differentiate between the contribution of machinery degradation and the degradation of the hull. A real-time assessment of the hull fouling condition will assist in the decision making process of when to clean the hull or dry dock, which is typically carried out at fixed intervals determined by the classification society for a specific type of vessel.

The proposed machine learning hull condition assessment takes advantage of the advancement of smart data collected when the vessel is in operation

(Lohr, 2012). Information gathered, such as vessel speed through the water, engine torque and thrust, and fuel consumption, is used as inputs and variables to learn and assess the hull condition. Data monitoring for ship performance analysis was first presented in 1989 (Chen, 1989) where data was saved on a ship database with manual input of the environmental condition every six hours and compared to the World Meteorological Organisation reports (Jensen et al., 1984). The advancement of technology has led to the possibility to record and transmit data to shore directly via satellite, whilst synchronised with up-to-date metocean datasets (Premkumar et al., 2000; Venkatesan et al., 2013).

Conventional use of shipboard measurements to estimate hull fouling is carried out by filtering the data using a mixture of computation simulation and empirical methods (Foteinos et al., 2017). The reliability of this methodology depends on the accuracy of each layer of filtering. A ship performance index using the Propulsion Diagnosis Number (PDno) was developed using real-time data and is claimed to address the hull and propeller roughness, however parameters such as propeller mean speed cannot be defined accurately (Deligiannis, 2017). Moreover, the method does not take into account the propeller cavitation and degradation as well as having constant fuel calorific values.

Data Analysis is improving our ability to understand complex phenomena much more rapidly than a priori physical models have done in the past (Anguita et al., 2010; Peng et al., 2010). Real-world systems are usually very complex and influenced by many exogenous factors, which make them very challenging to model, relying solely on the a priori knowledge of the problem (Witten et al., 2016). Data Analysis can use raw sensor data to provide useful information about the efficiency of a ship (Smith et al., 2013), identify operational profiles (Oneto et al., 2016), reduce the fuel consumption (Coraddu et al., 2015), and improve maintenance activities (Coraddu et al., 2016).

With the support of Data-Driven Models it is possible to exploit advanced statistical techniques to build models directly based on historical data produced and stored by the logging and monitoring equipment, without requiring any a priori knowledge of the underlining physical phenomena (Vapnik, 1998; Newton, 2000). This data represents strategic information for shipyards, operators, ship owners, and crews since it can be used for advisory, control, and fault detection purposes (Oneto et al., 2016). Furthermore, Data-Driven Models allow exploitation of exogenous data as well, such as

weather conditions, which could contain hidden information, potentially not easily representable with a conventional approach (Coraddu et al., 2015).

Frequently used Data Analysis techniques applied to different problems include Neural Networks (NNs), Kernel Methods (KMs), and Ensemble Methods (EMs). Examples of Data Analysis approaches applied to the marine industry can be found in (Palmé et al., 2011), where a standard Neural Network approach is used to improve monitoring of Gas Turbines, while Kernel based methods are applied in (Singer et al., 1995), and (Coraddu et al., 2016). In (Akinfiev et al., 2008) and (Bagavathiappan et al., 2013) image processing techniques are adopted for hull condition assessment. In (Basurko & Uriondo, 2015) the engine and propeller state is predicted adopting an Artificial Neural Network. A complete overview can be found in (Lazakis et al., 2016).

One of the most useful tools in real world data analytics problems is the novelty detection algorithm. Novelty (or outlier) detection methods address the problem of identifying new or unknown data that a data analytics system has not been trained with and was not previously aware of (Swersky et al., 2016; Markou & Singh, 2003; Pimentel et al., 2014). Novelty detection is also referred to as one-class classification because it is trained only on the one class of known data. Novelty detection is one of the fundamental problems in a classification system. A data analytics system cannot be trained with all of the possible object classes and hence the performance of the model will be poor for those classes that are under-represented in the training set. A good classification system must have the ability to differentiate between known and unknown objects during testing. Several models for novelty detection have been proposed, such as Support Vector Data Description, Gaussian Data Description, Parzen Window Data Description, k-Nearest Neighbor Data Description, and Global-Local Outlier Scores from Hierarchies (Markou & Singh, 2003; Pimentel et al., 2014).

The work presented in this paper is focused on building effective Data-Driven Models to predict the hull state condition utilising the data collected from an automation system when the vessel was in operation. The innovative contribution of this work lies in the fact that the proposed approach does not try to solve a traditional classification problem where the target is to estimate the label state of the hull. On the contrary, the problem has been tackled with an approach which, in principle, does not need any labeled samples since it searches for novel behaviour in the data through a novelty detection algorithm (Markou & Singh, 2003; Schölkopf et al., 2000). The

results show that with just a few samples from data collected during sea trials in the nominal state for the vessel, where the performances are unaffected by degradation due to hull and propeller fouling, it is possible to fine tune this methodology to achieve satisfactory performances in prediction of the hull state. Data obtained from Research Vessel (RV) The Princess Royal, a twin hull catamaran, operating mainly off the coast of Blyth, UK has been exploited to show the effectiveness of the proposed approaches and to benchmark them in a realistic maritime application.

The paper is structured as follows: in Section 2 a general explanation of the physical problem is provided, including the vessel description, the data acquisition system and data pre-processing. In Section 3 a description of the mathematical framework and the proposed ML techniques for the novelty detection are described. Section 4 illustrates the results obtained. Finally, in Section 5 the conclusions of the paper are drawn.

2. Vessel Description and Data available

2.1. Introduction

The novelty detection procedure proposed in this paper for the prediction of the hull state is applied to a case study based on the RV The Princess Royal. The vessel is a multi-purpose science and technology vessel with a flexible speed range for carrying out a wide variety of full-scale marine measurements and observations. The primary vessel duties include performance monitoring, coating/fouling inspection, cavitation and noise research. The main particulars of the vessel are described in Table 1, a complete description of the main vessel duties can be found in (Atlar et al., 2013).

Table 1: Main particulars of The Princess Royal

Ship feature	Value	Unit
Length overall	18.9	[m]
Length between Perpendicular	16.45	[m]
Breadth Moulded	7.03	[m]
Displacement (Lightship)	36.94	[t]
Draught (Lightship) (Amid - FP - AP)	1.65-1.6-1.7	[m]
Deadweight	7.32	[t]



Figure 1: RV The Princess Royal.

Hull surface degrades through time when a vessel is placed in the water. Figure 1 shows the two different hull conditions of the The Princess Royal in clean (a) and fouled (b) conditions. The proof of concept of the proposed method uses real time data collected when the vessel was in operation. Figures 2 and 3 illustrate samples of data for the problem description where the overlapping data sets of engine torque and fuel consumption from both port and starboard are presented. From Figures 4 and 5 it can be seen that the same overlapping is present when the engine speed is also considered. Finally in Figures 6 and 7 the fuel consumption is reported as a function of the vessel speed for both the shaft-lines. The figures show that for the two shaft-lines, port and starboard, a visual distinction is not possible between the clean and fouled hull conditions.

In order to classify and distinguish the hull conditions without visual identification, Data-Driven models are used to carry out hull condition assessment of the RV The Princess Royal. Physical parameters of the vessel and the particulars of the propulsion system are shown in Table 2.

2.2. Sea Trials

Two separate sea trials were carried out six months apart to allow fouling to take place on the hull surface under normal operational conditions. The first set of sea trials was conducted when the vessel was just out of dry dock and her hull was clean whilst the second set of sea trials was conducted when the vessel had been in operation for six months off the coast of Blyth, in the

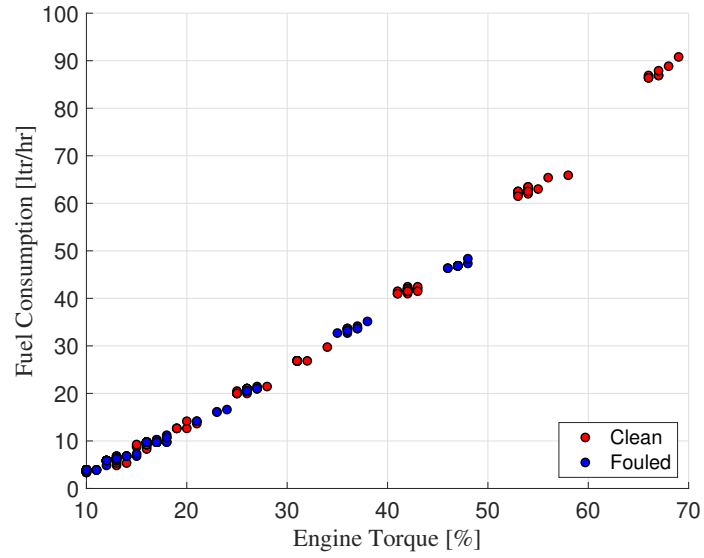


Figure 2: Overlapping data sets of fuel consumption and engine torque - starboard shaft

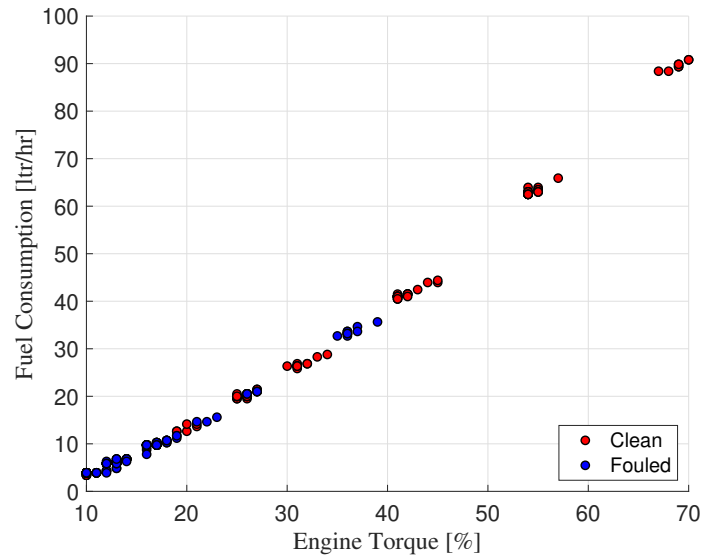


Figure 3: Overlapping data sets of fuel consumption and engine torque - port side shaft

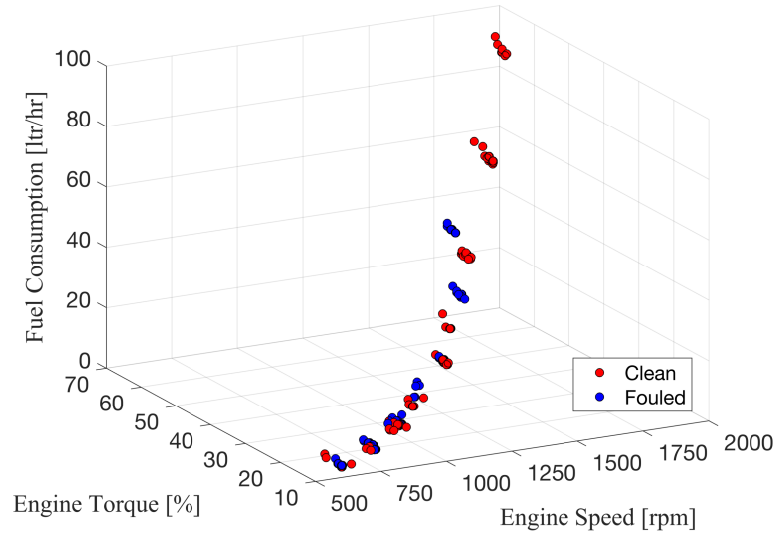


Figure 4: Overlapping data sets of fuel consumption, engine speed and engine torque - starboard shaft

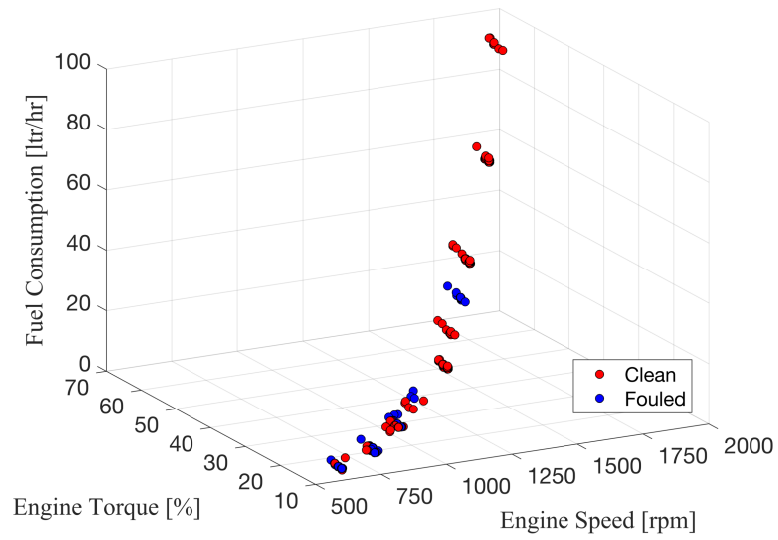


Figure 5: Overlapping data sets of fuel consumption, engine speed and engine torque - port side shaft

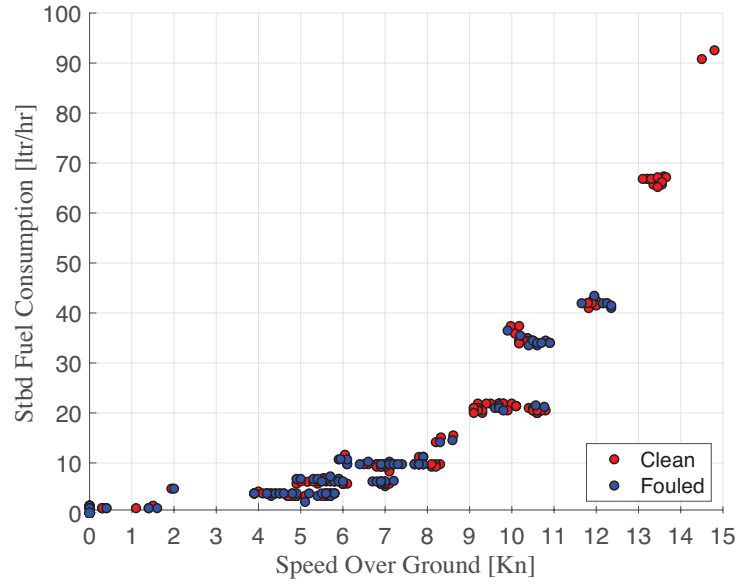


Figure 6: Overlapping data sets of fuel consumption and vessel speed - starboard shaft

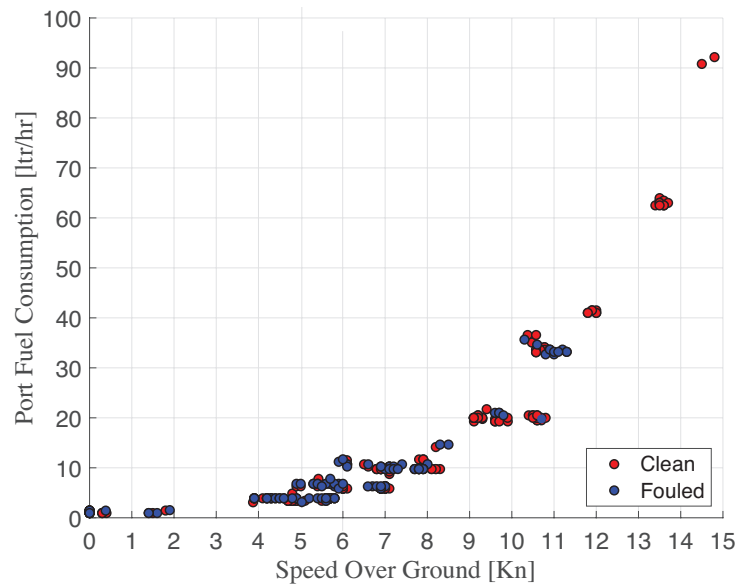


Figure 7: Overlapping data sets of fuel consumption and vessel speed - port side shaft

Table 2: Propeller, engine and gearbox particulars of The Princess Royal

Propeller particulars	
Number of Propellers	2
Propellers Type	Fixed
Propeller Diameter	0.75 [m]
Number of Blades	5
Engine particulars	
Number of Engines	2
Engine Make and Model	MAN D2676 LE443
Rated Power	537 [kW]
Rated Speed	2300 [rpm]
Rated Fuel Consumption	142 [Ltr/hr]
Gearbox particulars	
Number of Gearbox	2
Reduction Ratio	1.75:1

North Sea.

The sea trials are designed to allow the vessel to operate over the full range of allowable speeds. This is to allow the recording of all parameters that change with the speed of the vessel. During the sea trials, the ship track chosen covers the same ground area which is in line with ISO 15016:2015 (ISO 15016:2015). The vessel is allowed to run on a straight course for at least two miles between the same start and end points to ensure that the corresponding speed to the shaft power is achieved with a good accuracy. The sea trials were carried out at a range of engine rotational speeds which were 600, 750, 900, 1200, 1500 and 1750rpm. This was to capture the added wave resistance effect during higher speeds. Each speed run was carried out twice, once in the direction heading into the waves and once in the direction following the waves. The rudder was kept at a low angle to minimise its potential effect of vessel speed loss. The sea trials were conducted in a location of deep water with small variations of seabed contours to minimise any shallow water effects on the vessel and avoid the need to correct the data set. To compare data from the two sea trials, the weather condition needs to be taken into account. Experience of the vessel crew and weather forecasts were used to determine the calmest condition to carry out the sea trials to eliminate the effect of

further external variables such as wind, wave, and current that might affect or corrupt the data.

In order to further verify that the environmental conditions did not play a big role in this set of data, comparison was made with the weather data collected from the North East Coastal Observatory.

Given the location of the trials, data collected from the Newbiggin Ness Waverider buoy has been utilised. The buoy provides information about significant wave height, wind speed and peak period of the spectrum. A detailed list of parameters is reported in Table 3.

Table 3: Newbiggin Ness buoy data

Parameter	Unit
Significant Wave Height	[m]
Wave zero up crossing period	[s]
Wind Wave Significant Wave Height	[m]
Wind Wave Peak Period	[s]
Wind Wave Peak Period Direction	[°]
Swell Significant Wave Height	[m]
Swell Peak Period	[s]
Swell Peak Period Direction	[°]
Sea Surface Elevation	[m]
Sea Temperature	[°C]

Influence from the environmental conditions affects the power required to keep the vessel on course.

The use of the rudder to maintain the vessel heading will increase resistance and causes changes in the parameters recorded such as shaft torque, shaft speed and ship speed.

Beaufort number 3 was recorded during the dirty hull sea trial whilst Beaufort number 4 was recorded on the day of the clean hull sea trial.

This is in accordance with the International Towing Tank Conference (ITTC) recommendation that a vessel less than 100 m in length should conduct sea trials at wind speeds not higher than Beaufort number 5.

In terms of the sea state, the total wave height was calculated from the recorded significant wave heights and swells during the sea trials. The total wave height was 0.77 m for the clean hull sea trials and 0.35 m for the dirty hull sea trials. The sea state conditions also satisfy the ITTC recommendation where the total wave height calculated from the spectrum H is within the limit set:

$$H \leq 2.25 \sqrt{L_{PP}/100} \quad (1)$$

2.3. Data Logging System

The data recorded on the vessel uses a single common platform installed on-board RV The Princess Royal to enable-real time data to be collected and time stamped accurately.

The data gathered includes the monitoring systems of fuel oil and engine monitoring.

Strain gauges are fitted on the shaft through an in-house design which captures the thrust, torque and bending moment of the shaft on both port and starboard sides. The data collected from this single platform monitoring system has been verified through other studies that monitor vessel performance and holistic energy flow on the vessel (Carchen et al., 2017). Moreover, data from the automatic identification system (AIS) such as vessel speed over ground, speed through water, vessel position (longitude and latitude), course over ground and vessel true heading are also recorded. Weather data, wind apparent speed and apparent direction are obtained through the NMEA0183 specification which includes data from an anemometer and gyro-compass. The wind apparent speed values take into account the vessel heading and the values are important to quantify additional wind load applied on the vessel at different speeds.

Data from the built-in sensors of the engines was obtained through a CANBus data source. Table 4 shows the measured values available from the sea trials measurements.

2.4. Data Pre-Processing

Prior to building the machine learning models, it is necessary to perform data-cleaning steps to prepare the data for training and testing. The quality of the prepared data can significantly influence the success of a model (Kotsiantis et al., 2007). Since all data collected was from physical systems, outliers with extreme values were removed from the dataset.

By performing preliminary analysis on the remaining data, useful information has been extracted. Due to the nature of the sea-trials, there were some factors that had to be filtered out before the classification process. The factors that were not considered for the classification process included the geographical coordinates (Features 3 and 4) that map the position of the vessel

Table 4: Measured values available from the monitoring system

Feature	Variable name	Unit
x_1	Magnetic Heading	[deg]
x_2	True Heading	[deg]
x_3	Latitude	[$^{\circ}$, ', '']
x_4	Longitude	[$^{\circ}$, ', '']
x_5	Speed through water	[kn]
x_6	Water Depth	[m]
x_7	Humidity	[%]
x_8	Rudder angle	[deg]
x_9	Course over ground	[deg]
x_{10}	Wind Apparent Speed	[m/s]
x_{11}	Wind Apparent Direction	[deg]
x_{12}	Air Temperature	[$^{\circ}C$]
x_{13}	Air Pressure	[mbars]
x_{14}	Relative Humidity	[%]
x_{15}	Coolant Pressure (port)	[bar]
x_{16}	Coolant Temperature (port)	[$^{\circ}C$]
x_{17}	Engine Speed (port)	[rpm]
x_{18}	Engine Torque (port)	[%]
x_{19}	Fuel Delivery Pressure (port)	[bar]
x_{20}	Oil Pressure (port)	[bar]
x_{21}	Oil Temperature (port)	[$^{\circ}C$]
x_{22}	Crankcase Pressure (port)	[bar]
x_{23}	Oil Level (port)	[%]
x_{24}	Fuel Flow (port)	[Ltr/hr]
x_{25}	Fuel Return (port)	[%]
x_{26}	Fuel Supply Pressure (port)	[Ltr]
x_{27}	Fuel Consumption (port)	[Ltr/hr]
x_{28}	Coolant Pressure (starboard)	[bar]
x_{29}	Coolant Temperature (starboard)	[$^{\circ}C$]
x_{30}	Engine Speed (starboard)	[rpm]
x_{31}	Engine Torque (starboard)	[%]
x_{32}	Fuel Delivery Pressure (starboard)	[bar]
x_{33}	Oil Pressure (starboard)	[bar]
x_{34}	Oil Temperature (starboard)	[$^{\circ}C$]
x_{35}	Crankcase Pressure (starboard)	[bar]
x_{36}	Oil Level (starboard)	[%]
x_{37}	Fuel Flow (starboard)	[Ltr/hr]
x_{38}	Fuel Return (starboard)	[%]
x_{39}	Fuel Supply Pressure (starboard)	[bar]
x_{40}	Fuel Consumption (starboard)	[Ltr/hr]

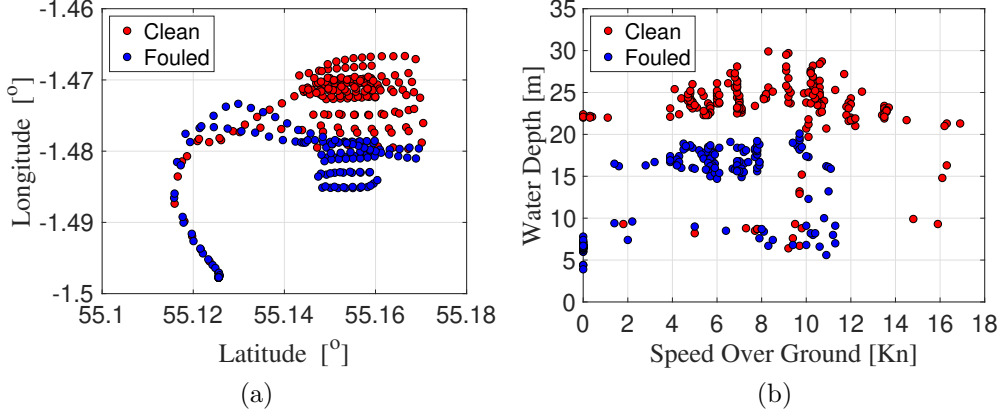


Figure 8: Trials geographical coordinates and water depth

during sea trials with corresponding water depth (Feature 5). These parameters have clear distinction shown in Figure 8 and will have a direct impact if taken into consideration as a dataset for the machine learning, creating a false indication of hull condition.

Figure 9 shows the **Principal Component Analysis** of the data which is possibly the best visual understanding and a useful tool to understanding the complexity of this classification problem. In Figure 9 two sets of data are presented. The first set is characterised by the fouled hull condition (blue coloured), while the second set of data (red coloured) is characterised by the clean hull condition. The figure shows, in the Euclidean plane, that the two sets of data are not linearly separable.

3. Novelty detection

In this section, the authors will present the Machine Learning (ML) techniques adopted to assess the hull condition of the vessel described in Section 2, based on the data outlined in the same section.

Consider an input space $\mathcal{X} \subseteq \mathbb{R}^d$ and an output space \mathcal{Y} . Note that, for the purpose of this paper, \mathcal{X} takes into account the different sensors' measurements, also called features, reported in Table 4, while the output space $\mathcal{Y} \in \{\pm 1\}$ represents a clean hull with -1 and a fouled hull with $+1$ (see Section 2). ML techniques aim at estimating the unknown rule $\mu : \mathcal{X} \rightarrow \mathcal{Y}$ which associates an element $y \in \mathcal{Y}$ to an element $x \in \mathcal{X}$. Note

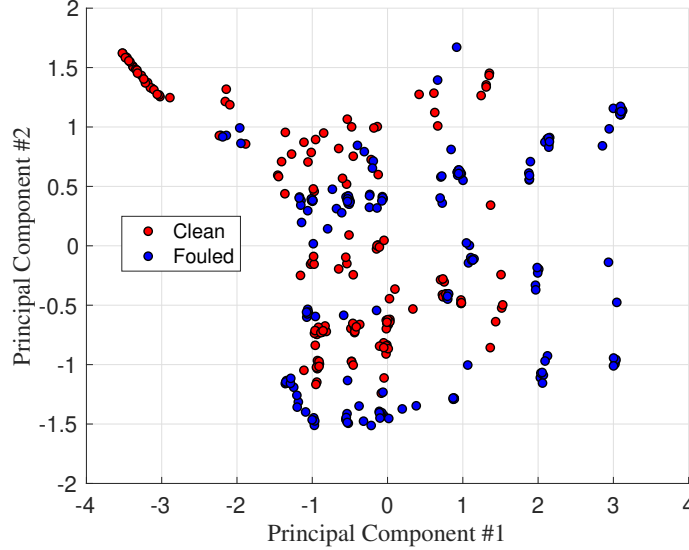


Figure 9: Principal Component Analysis - 2 main components

that, in general, μ can be non-deterministic. An ML technique estimates μ through a learning algorithm $\mathcal{A}_{\mathcal{H}} : \mathcal{D}_n \times \mathcal{F} \rightarrow h$, characterised by its set of hyperparameters \mathcal{H} , which maps a series of examples of the input/output relation contained in a dataset of n samples $\mathcal{D}_n : \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$ into a function $f : \mathcal{X} \rightarrow \mathcal{Y}$ chosen in a set of possible ones \mathcal{F} .

When both \mathbf{x}_i and y_i with $i \in \{1, \dots, n\}$ are available, the problem is termed supervised and consequently supervised ML techniques are adopted (Vapnik, 1998). When only \mathbf{x}_i with $i \in \{1, \dots, n\}$ are available, which means that the associated element of the output space y_i with $i \in \{1, \dots, n\}$ is not explicitly known, it has to be assumed that “similar” \mathbf{x}_i values are associated with “similar” y_i where the concept of similarity is something that needs to be defined based on μ . In the latter case, the ML problems are called unsupervised, and consequently, unsupervised ML techniques need to be adopted (Hastie et al., 2009).

Anomaly (novelty or outlier) detection is a common example of an unsupervised learning problem where the unknown $y \in \mathcal{Y}$ can assume only two possible values: -1 for “non-anomaly” and $+1$ for “anomaly” (Shawe-Taylor & Cristianini, 2004). In the proposed framework anomaly means that the ship is fouled.

The error that f generates in approximating μ is measured with reference to a loss function $\ell : \mathcal{X} \times \mathcal{Y} \times \mathcal{F} \rightarrow [0, \infty)$. Obviously, the error that f generates over \mathcal{D}_n , is optimistically biased since \mathcal{D}_n has been used, together with \mathcal{F} , for building f itself. For this reason, another set of fresh data, composed of m samples and called the test set $\mathcal{T}_m = \{(\mathbf{x}_1^t, y_1^t), \dots, (\mathbf{x}_m^t, y_m^t)\}$, needs to be used. Note that, $\mathbf{x}_i^t \in \mathcal{X}$ and $y_i^t \in \mathcal{Y}$ with $i \in \{1, \dots, m\}$, and the association of y_i^t to \mathbf{x}_i^t is again made based on μ . Moreover, both for supervised and unsupervised problems \mathcal{T}_m must contain both $\mathbf{x}_i^t \in \mathcal{X}$ and $y_i^t \in \mathcal{Y}$ with $i \in \{1, \dots, m\}$ to estimate the error of f , while, for unsupervised learning problems, y_i with $i \in \{1, \dots, n\}$ in \mathcal{D}_n is unknown.

3.1. Measuring the Error

In order to perform this analysis, different measures of error must be defined, also called **indices of performance**, which are able to characterise the quality of the hull condition prediction system. Once f has been chosen based on \mathcal{D}_n it is possible to use the new set of data \mathcal{T}_m in order to compute the error based on different losses. The choice of the loss strongly depends on the problem under examination (Rosasco et al., 2004).

In the anomaly detection framework, the most natural choice as loss function is the Hard loss one, which counts the number of misclassified samples $\ell_H(f(\mathbf{x}), y) = [f(\mathbf{x}) \neq y]$. Note that the Iverson bracket notation is used. Based on the Hard loss it is possible to define different indexes of performance (Powers, 2011):

- the Average Misclassifications Rate (AMR) is the mean number of misclassified samples: $\text{AMR} = \frac{1}{m} \sum_{i=1}^m \ell_H(f(\mathbf{x}_i^t), y_i^t)$;
- the Confusion Matrix, which measures four different quantities:
 - $\text{TN} = 100/m \sum_{i=1}^m [f(\mathbf{x}_i^t) = y_i^t \wedge y_i^t = -1]$ which is the percentage of true negative;
 - $\text{TP} = 100/m \sum_{i=1}^m [f(\mathbf{x}_i^t) = y_i^t \wedge y_i^t = +1]$ which is the percentage of true positive;
 - $\text{FN} = 100/m \sum_{i=1}^m [f(\mathbf{x}_i^t) \neq y_i^t \wedge y_i^t = -1]$ which is the percentage of false negative;
 - $\text{FP} = 100/m \sum_{i=1}^m [f(\mathbf{x}_i^t) \neq y_i^t \wedge y_i^t = +1]$ which is the percentage of false positive.

3.2. Machine Learning Techniques

In this section, the unsupervised learning algorithms used for building the data driven models are presented. In addition, how to tune the model

performances by tuning their hyperparameters during the so-called Model Selection (MS) phase (Kohavi et al., 1995; Bartlett et al., 2002; Anguita et al., 2012) will also be discussed. Finally, checks for possible spurious correlation in the data will be carried out by performing the Feature Selection (FS) phase (Guyon & Elisseeff, 2003; Friedman et al., 2001; Chang & Lin, 2008; Yoon et al., 2005; Hong, 1997; Franceschi et al., 2017). In fact, once f has been built based on the different learning algorithm and has been confirmed to be a sufficiently accurate representation of μ , it can be interesting to investigate how the model f is affected by the different features that have been exploited to build f itself during the feature ranking procedure (Guyon & Elisseeff, 2003). For some algorithms, the feature ranking procedure is a by-product of the learning process itself, allowing simple checking of the physical plausibility of f .

3.2.1. *Unsupervised Learning Algorithms for Anomaly Detection*

Unsupervised ML methods can be divided into different families. Since this work deals only with anomaly detection problems, **the authors** considered only the most well-known and effective techniques for solving these problems according to (Swersky et al., 2016). In particular (Swersky et al., 2016) shows that two anomaly detection methods based on Support Vector Machines (SVM) (Cristianini & Shawe-Taylor, 2000) and k-nearest neighbourhood (KNN) (Cover & Hart, 1967) respectively, are the most commonly used and reliable techniques in this context.

In particular One-Class SVM (OCSVM) is a boundary-based anomaly detection method, inspired by SVM, which encloses the inlier class in a minimum volume hypersphere by minimizing a Tikhonov regularization problem, similar to the one reported for the SVM framework. Like traditional SVMs, OCSVM can also be extended to non-linearly transformed spaces using the “Kernel trick” for distances. The hyperparameters OCSVM $\mathcal{H}^{\text{OCSVM}}$ are: the kernel, which is usually fixed and in this paper the authors chose the Gaussian Kernel for the reasons described in (Keerthi & Lin, 2003; Oneto et al., 2015), its hyperparameter h_1 and the regularization hyperparameter h_2 .

The Global KNN (GKNN), inspired by the KNN, was originally introduced as an unsupervised distance-based outlier detection method (Ramswamy et al., 2000; Swersky et al., 2016). The hyperparameter GKNN $\mathcal{H}^{\text{GKNN}}$ is the number of neighbours to be considered h_1 .

3.2.2. Model Selection

MS deals with the problem of tuning the hyperparameters of each learning algorithm (Anguita et al., 2012). Several methods exist for MS including resampling methods, like the well-known k -Fold Cross Validation (Kohavi et al., 1995) or the nonparametric Bootstrap (BTS) approach (Efron & Tibshirani, 1994; Anguita et al., 2000), which represent the state-of-the-art MS approaches when targeting real-world applications. Resampling methods rely on a simple idea: the original dataset \mathcal{D}_n is resampled once or many (n_r) times, with or without replacement, to build two independent datasets called training, and validation sets, \mathcal{L}_l^r and \mathcal{V}_v^r respectively, with $r \in \{1, \dots, n_r\}$. Note that $\mathcal{L}_l^r \cap \mathcal{V}_v^r = \emptyset$, $\mathcal{L}_l^r \cup \mathcal{V}_v^r = \mathcal{D}_n$. Then, in order to select the best combination of hyperparameters \mathcal{H} in a set of possible ones $\mathfrak{H} = \{\mathcal{H}_1, \mathcal{H}_2, \dots\}$ for the algorithm $\mathcal{A}_{\mathcal{H}}$ or, in other words, to perform the MS phase, the following procedure has to be applied:

$$\mathcal{H}^* : \min_{\mathcal{H} \in \mathfrak{H}} \frac{1}{n_r} \sum_{r=1}^{n_r} \frac{1}{v} \sum_{(\mathbf{x}_i, y_i) \in \mathcal{V}_v^r} \ell(\mathcal{A}_{\mathcal{H}, \mathcal{L}_l^r}(\mathbf{x}_i), y_i), \quad (2)$$

where $\mathcal{A}_{\mathcal{H}, \mathcal{L}_l^r}$ is a model built with the algorithm \mathcal{A} with its set of hyperparameters \mathcal{H} and with the data \mathcal{L}_l^r . Since the data in \mathcal{L}_l^r is independent from that in \mathcal{V}_v^r , the idea is that \mathcal{H}^* should be the set of hyperparameters which achieves a small error on a data set that is independent from the training set.

Note that, for the anomaly detection problem, the algorithms do not need any labels in \mathcal{L}_l^r , consequently only labeled data for \mathcal{V}_v^r is required (Swersky et al., 2016).

In this paper the BTS is used because it represents the state-of-the-art approach (Efron & Tibshirani, 1994; Anguita et al., 2012). For implementing the bootstrap, $l = n$ and \mathcal{L}_l^r must be sampled with replacement from \mathcal{D}_n , while \mathcal{V}_v^r and \mathcal{T}_t^r are sampled without replacement from the sample of \mathcal{D}_n that has not been sampled in \mathcal{L}_l^r (Efron & Tibshirani, 1994; Anguita et al., 2012). Note that for the bootstrap procedure $r \leq \binom{2n-1}{n}$.

3.3. Feature Selection

Once the models are built and have been confirmed to be a sufficiently accurate representation of the hull condition, how these models are affected by the different features used in the model identification phase (see Table 4) **is investigated**.

In DA this procedure is called Feature Selection (FS) or Feature Ranking (Guyon & Elisseeff, 2003; Friedman et al., 2001; Chang & Lin, 2008; Yoon et al., 2005; Hong, 1997). This process allows investigating whether the importance of those features, that are known to be relevant from a physical perspective, is appropriately described by the different models. The failure of the statistical model to properly account for the relevant features might indicate poor quality in the measurements or spurious correlations. FS therefore represents an important step of model verification, since it should generate consistent results with the available knowledge of the physical system under examination.

Given the limited number of features available in this work, the authors decided to proceed with a brute force approach (Guyon & Elisseeff, 2003). In particular, one feature at a time was removed from the training phase, and a check on the effect of this removal on the final mean AMR of the model, averaged over multiple runs of the learning phase, was performed.

The higher the mean increase in AMR due to the removal, the more important the removed feature. This procedure was adopted since it is relatively straightforward to implement during the main prediction process and computationally inexpensive.

4. Experimental results

In this section, the results obtained from the different novelty detection methods are discussed. The problem was solved in an unsupervised fashion by modelling the problem as a novelty detection one to further reduce the need for labeled data. In the unsupervised case, different dimensions of the training set and the MS procedure were performed as described in Section 3.2.2:

1. OCSVM: the set of hyperparameters is $\mathcal{H}^{OCSVM} = \{h_1, h_2\}$ and authors chose it in $\mathfrak{H}^{OCSVM} = \{10^{-4}, 10^{-3.7}, \dots, 10^3\} \times \{10^{-4}, 10^{-3.8}, \dots, 10^{-1.0}\}$;
2. GKNN: the set of hyperparameters is $\mathcal{H}^{GKNN} = \{h_1\}$ and authors chose it in $\mathfrak{H}^{GKNN} = \{1, 3, 7, 13, 27, 51\}$;

The \mathcal{V}_v^r cardinality was varied, to test the possibility of building an efficient model with a small number of samples.

The performance of each model was measured according to the metrics described in Section 3. Each experiment was performed 30 times in order to

	v	AMR	TP	TN	FP	FN
OCSVM	10	0.04 ± 0.01	47.8 ± 1.2	47.8 ± 1.0	2.2 ± 1.0	1.8 ± 1.1
	20	0.04 ± 0.01	48.0 ± 0.9	48.0 ± 1.1	2.0 ± 1.0	2.0 ± 1.0
	30	0.03 ± 0.01	48.4 ± 1.0	48.4 ± 0.9	1.6 ± 1.0	1.4 ± 1.0
GKNN	10	0.05 ± 0.02	47.6 ± 2.3	47.6 ± 1.8	2.4 ± 1.9	2.6 ± 1.9
	20	0.04 ± 0.02	48.2 ± 2.0	48.2 ± 1.9	1.8 ± 2.3	2.2 ± 2.0
	30	0.03 ± 0.01	48.7 ± 1.2	48.7 ± 1.1	1.3 ± 1.0	1.7 ± 0.8

Table 5: AMR, TP, TN, FP, and FN of the models learned with the different algorithms (OCSVM and GKNN) when $l = 150$ and $v \in \{10, 20, 30\}$.

obtain statistically relevant results, and the t-student 95% confidence interval is reported.

For OCSVM the R package of Meyer et al. (2015) was used to test the OCSVM while GKNN was implemented by the authors in R.

The AMR, the TP, the TN, the FP, and the FN of the models learned with the different algorithms (OCSVM and GKNN) are reported in Table 5, where the number of unlabelled samples in the learning set is $l \in \{150\}$ and when varying the number of labeled samples in the validation set $v \in \{10, 20, 30\}$ (half positively and half negatively labeled). In Table 6 the AMR, the TP, the TN, the FP, and the FN of the models learned with the different algorithms (OCSVM and GKNN) are reported, where the number of labeled samples in the validation set is $v \in \{30\}$ and when varying the number of unlabelled samples in the learning set $l \in \{30, 70, 150\}$ (half positively and half negatively labeled).

From the tables it is possible to observe that:

- both OCSVM and GKNN perform quite well on the problem and there is no clear distinction between them;
- changing the l or v values does not significantly affect the performance of the models;
- with just a few labeled samples, around 10, it is possible to obtain satisfactory accuracy; this is quite a remarkable result, since 10 samples can be easily manually labeled by an expert operator;
- FP and FN rates are quite balanced and this is a further indication of the quality of the result.

Finally, in Table 7 the FS phase is presented. In particular, for each

	l	AMR	TP	TN	FP	FN
OCSVM	30	0.22 ± 0.11	39.5 ± 10.8	39.5 ± 11.3	10.5 ± 10.7	11.5 ± 11.5
	70	0.07 ± 0.04	46.4 ± 3.5	46.4 ± 3.6	3.6 ± 3.7	3.4 ± 3.8
	150	0.03 ± 0.01	48.4 ± 1.0	48.4 ± 0.9	1.6 ± 1.0	1.4 ± 1.0
GKNN	30	0.25 ± 0.15	37.9 ± 15.0	37.9 ± 10.5	12.1 ± 14.3	12.9 ± 16.9
	70	0.11 ± 0.06	45.1 ± 6.6	45.1 ± 6.2	4.9 ± 6.0	6.1 ± 6.1
	150	0.03 ± 0.01	48.7 ± 1.2	48.7 ± 1.1	1.3 ± 1.0	1.7 ± 0.8

Table 6: AMR, TP, TN, FP, and FN of the models learned with the different algorithms (OCSVM and GKNN) when $v = 30$ and $l \in \{30, 70, 150\}$.

feature (see Table 4) the top five sorted list of features (from the most important one to the least important) together with its mean increase in AMR (as described in Section 3.3) is reported. The results of both OCSVM and GKNN are presented with $l = 150$ and $v = 30$.

From Table 7 it is possible to observe that both the OCSVM and GKNN models can adequately account for the relevant features as the outcome is consistent with the available knowledge. As expected, the parameters such as speed through the water, shaft torques and shaft speeds have the highest predictive power. In fact, these parameters are the most important for the fuel consumption assessment as reported in (Coraddu et al., 2015), and in turn for the hull and propeller fouling estimation.

These results also support the assumption made by the authors in Section 2.2: that the meteo-marine conditions do not affect the model for this particular case. The authors conclude that the added resistance in waves and the wind resistance contributions are negligible compared to the calm-water resistance, and under these environmental conditions, the proposed novelty detection is a reliable tool for the hull and propeller fouling assessment.

OCSVM		
x_5	Speed through water [kn]	0.12 ± 0.03
x_{17}	Engine Speed (port) [rpm]	0.07 ± 0.03
x_{30}	Engine Speed (starboard) [rpm]	0.07 ± 0.03
x_{18}	Engine Torque (port) [%]	0.06 ± 0.04
x_{31}	Engine Torque (starboard) [%]	0.06 ± 0.04

GKNN		
x_5	Speed through water [kn]	0.14 ± 0.04
x_{17}	Engine Speed (port) [rpm]	0.08 ± 0.03
x_{18}	Engine Torque (port) [%]	0.07 ± 0.04
x_{30}	Engine Speed (starboard) [rpm]	0.07 ± 0.04
x_{31}	Engine Torque (starboard) [%]	0.06 ± 0.04

Table 7: The top five sorted list of features (from the most important one to the least important) together with its mean increase in AMR (as described in Section 3.3) computed with both OCSVM and GKNN with $l = 150$ and $v = 30$.

5. Conclusions

In this paper, the authors presented two Data Analysis unsupervised techniques for the identification of the most influential parameters for the prediction of hull and propeller fouling in real operations based on data measured by the onboard automation systems. The preferable time interval for cleaning the ship’s hull depends on many factors such as the type of paint covering the underwater surface, the ratio between the standing and movement times, the water temperature and surrounding environmental condition of the hull. The proposed approach allows for better management of assets which in turn allows owners to predict the hull condition and suggest the best time for dry docking and carrying out hull maintenance work. Moreover, the method can be applied in many ship operational activities: from the assessment of the right intervals between maintenance actions for the propeller and the hull cleaning to the efficiency of these measures.

The Data-Driven approach proved to be successful in the identification of the fouled hull condition and is validated as a tool for a better understanding

of the extent of degradation of the hull and propeller. The novelty detection method shows a satisfactory accuracy and this is a remarkable result, since the few (10) required samples can be easily manually labeled by an operator. This study proved that it is possible to treat the problem of the prediction of the hull condition in an unsupervised fashion. The proposed models can be adopted for real-time applications directly onboard, to easily and quickly identify maintenance **requirements** and limit any decrease of performances. Although it was not tested in this particular study, the method can be foreseen to be of use also for the evaluation of energy-saving technologies, such as new propeller designs and sails. The authors recognise the contribution of this work as a significant development for speeding up the uptake of new technologies, particularly for providing a better way to estimate performance and compare it to a reliable benchmark.

References

- Akinfiev, T. S., Armada, M. A., & Fernandez, R. (2008). Nondestructive testing of the state of a ship's hull with an underwater robot. *Russian Journal of Nondestructive Testing*, 44, 626–633.
- Anguita, D., Boni, A., & Ridella, S. (2000). Evaluating the generalization ability of support vector machines through the bootstrap. *Neural Processing Letters*, 11, 51–58.
- Anguita, D., Ghio, A., Greco, N., Oneto, L., & Ridella, S. (2010). Model selection for support vector machines: Advantages and disadvantages of the machine learning theory. In *International Joint Conference on Neural Networks*.
- Anguita, D., Ghio, A., Oneto, L., & Ridella, S. (2012). In-sample and out-of-sample model selection and error estimation for support vector machines. *IEEE Transactions on Neural Networks and Learning Systems*, 23, 1390–1406.
- Atlar, M., Aktas, B., Sampson, R., Seo, K., Viola, I., Fitzsimmons, P., & Fetherstonhaugh, C. (2013). A multi-purpose marine science & technology research vessel for full-scale observations and measurements. *AMT'13, Gdansk*, .

- Bagavathiappan, S., Lahiri, B. B., Saravanan, T., Philip, J., & Jayakumar, T. (2013). Infrared thermography for condition monitoring-a review. *Infrared Physics & Technology*, 60, 35–55.
- Bartlett, P. L., Boucheron, S., & Lugosi, G. (2002). Model selection and error estimation. *Machine Learning*, 48, 85–113.
- Basurko, O. C., & Uriondo, Z. (2015). Condition-based maintenance for medium speed diesel engines used in vessels in operation. *Applied Thermal Engineering*, 80, 404–412.
- Carchen, A., Pazouki, K., & Atlar, M. (2017). Development of an online ship performance monitoring system dedicated for biofouling and anti-fouling coating analysis. In *Hull Performance and Insight Conference (HullPIC)*.
- Chang, Y. W., & Lin, C. J. (2008). Feature ranking using linear svm. In *WCCI Causation and Prediction Challenge*.
- Chen, H. (1989). A new approach to ship speed performance monitoring and prediction. In *Ship technology and Research Star Symposium*. volume 14.
- Coraddu, A., Oneto, L., Baldi, F., & Anguita, D. (2015). Ship efficiency forecast based on sensors data collection: Improving numerical models through data analytics. In *OCEANS 2015-Genova*.
- Coraddu, A., Oneto, L., Ghio, A., Savio, S., Anguita, D., & Figari, M. (2016). Machine learning approaches for improving condition-based maintenance of naval propulsion plants. *Proceedings of the Institution of Mechanical Engineers Part M: Journal of Engineering for the Maritime Environment*, 230, 136–153.
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13, 21–27.
- Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.
- Deligiannis, P. (2017). Ship performance indicator. *Marine Policy*, 75, 204 – 209.

- Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.
- Flemming, H. (2002). Biofouling in water systems—cases, causes and countermeasures. *Applied microbiology and biotechnology*, 59, 629–640.
- Foteinos, M., Tzanos, E., & Kyrtatos, N. (2017). Ship hull fouling estimation using shipboard measurements, models for resistance components, and shaft torque calculation using engine model. *Journal of Ship Research*, 61, 64–74.
- Franceschi, L., Donini, M., Frasconi, P., & Pontil, M. (2017). On hyperparameter optimization in learning systems, .
- Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning*. Springer series in statistics Springer, Berlin.
- Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *The Journal of Machine Learning Research*, 3, 1157–1182.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). Unsupervised learning. In *The elements of statistical learning*.
- Hong, S. J. (1997). Use of contextual information for feature ranking and discretization. *IEEE Transactions on Knowledge and Data Engineering*, 9, 718–730.
- IMO (2011). *A transparent and reliable hull and propeller performance standard, MEPC 63/4/8*. Technical Report International Maritime Organisation, Marine Environment Protection Committee (MEPC).
- ISO 15016:2015 (2015). *Ships and marine technology — Guidelines for the assessment of speed and power performance by analysis of speed trial data*. Standard International Organization for Standardization Geneva, CH.
- Jensen, N. O., Petersen, E. L., & Troen, L. (1984). Extrapolation of mean wind statistics with special regard to wind energy applications. *World Climate Programme Report No. WCP-86*, .
- Kane, D. (2012). *Marine vessel environmental performance (MVEP) assessment guide. Energy efficiency: Hull and propeller operations and maintenance*.

- Keerthi, S. S., & Lin, C. J. (2003). Asymptotic behaviors of support vector machines with gaussian kernel. *Neural computation*, 15, 1667–1689.
- Kohavi, R. et al. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International Joint Conference on Artificial Intelligence*.
- Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2007). Data preprocessing for supervised learning. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 1, 4091–4096.
- Lazakis, I., Dikis, K., Michala, A. L., & Theotokatos, G. (2016). Advanced ship systems condition monitoring for enhanced inspection, maintenance and decision making in ship operations. *Transportation Research Procedia*, 14, 1679–1688.
- Lohr, S. (2012). The age of big data. *New York Times*, 11.
- Markou, M., & Singh, S. (2003). Novelty detection: a review. *Signal processing*, 83, 2481–2497.
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2015). e1071: Misc functions of the department of statistics, probability theory group (formerly: E1071), tu wien. R package. R package version 1.6-7.
- Newton, L. R. (2000). Data-logging in practical science: research and reality. *International Journal of Science Education*, 22, 1247–1259.
- Oneto, L., Anguita, D., Coraddu, A., Cleophas, T., & Xepapa, K. (2016). Vessel monitoring and design in industry 4.0: A data driven perspective. In *IEEE International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow*.
- Oneto, L., Ghio, A., Ridella, S., & Anguita, D. (2015). Support vector machines and strictly positive definite kernel: The regularization hyperparameter is more important than the kernel hyperparameters. In *IEEE International Joint Conference on Neural Networks*.
- Palmé, T., Breuhaus, P., Assadi, M., Klein, A., & Kim, M. (2011). New alstom monitoring tools leveraging artificial neural network technologies. In *Turbo Expo: Turbine Technical Conference and Exposition*.

- Peng, Y., Dong, M., & Zuo, M. J. (2010). Current status of machine prognostics in condition-based maintenance: a review. *The International Journal of Advanced Manufacturing Technology*, 50, 297–313.
- Pimentel, M. A. F., Clifton, C. A., Clifton, L., & Tarassenko, L. (2014). A review of novelty detection. *Signal Processing*, 99, 215–249.
- Powers, D. M. (2011). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2, 37–63.
- Premkumar, K., Ravichandran, M., Kalsi, S., Sengupta, D., & Gadgil, S. (2000). First results from a new observational system over the indian seas. *Current Science*, 78, 323–330.
- Ramaswamy, S., Rastogi, R., & Shim, K. (2000). Efficient algorithms for mining outliers from large data sets. In *ACM Sigmod Record*.
- Rosasco, L., De Vito, E., Caponnetto, A., Piana, M., & Verri, A. (2004). Are loss functions all the same? *Neural Computation*, 16, 1063–1076.
- Schölkopf, B., Williamson, R. C., Smola, A. J., Shawe-Taylor, J., & Platt, J. C. (2000). Support vector method for novelty detection. In *Advances in neural information processing systems* (pp. 582–588).
- Schultz, M., Bendick, J., Holm, E., & Hertel, W. (2011). Economic impact of biofouling on a naval surface ship. *Biofouling*, 27, 87–98.
- Schultz, M. P. (2007). Effects of coating roughness and biofouling on ship resistance and powering. *Biofouling*, 23, 331–341.
- Shawe-Taylor, J., & Cristianini, N. (2004). *Kernel methods for pattern analysis*. Cambridge university press.
- Singer, R. M., Gross, K. C., & King, R. W. (1995). A pattern-recognition-based, fault-tolerant monitoring and diagnostic technique. In *Argonne National Lab., Idaho Falls, ID (United States)*.
- Smith, T., O’Keeffe, E., Aldous, L., & Agnolucci, P. (2013). Assessment of shipping’s efficiency using satellite ais data. In *International Council on Clean Transportation*.

- Soares, C. G., Garbatov, Y., Zayed, A., & Wang, G. (2009). Influence of environmental factors on corrosion of ship structures in marine atmosphere. *Corrosion Science*, 51, 2014–2026.
- Swersky, L., Marques, H. O., Sander, J., Campello, R. J., & Zimek, A. (2016). On the evaluation of outlier detection and one-class classification methods. In *IEEE International Conference on Data Science and Advanced Analytics*.
- Tupper, E. C. (2013). *Introduction to naval architecture*. Butterworth-Heinemann.
- Vapnik, V. N. (1998). *Statistical learning theory*. Wiley New York.
- Venkatesan, R., Shamji, V., Latha, G., Mathew, S., Rao, R., Muthiah, A., & Atmanand, M. (2013). In situ ocean subsurface time-series measurements from omni buoy network in the bay of bengal. *Current Science*, (pp. 1166–1177).
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
- Yoon, H., Yang, K., & Shahabi, C. (2005). Feature subset selection and feature ranking for multivariate time series. *IEEE transactions on knowledge and data engineering*, 17, 1186–1198.